

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 16:34:37 : Search time 23.7857 seconds  
(without alignments)  
28.011 Million cell updates/sec

Title: us-09-856-070-18

Perfect score: 24

Sequence: 1 KEELM 5

Scoring table: HLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2500000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_geneset101002.*				
1:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*			
2:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*			
3:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*			
4:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*			
5:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*			
6:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*			
7:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*			
8:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*			
9:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*			
10:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*			
11:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*			
12:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*			
13:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*			
14:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*			
15:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*			
16:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*			
17:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*			
18:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*			
19:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*			
20:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*			
21:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*			
22:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*			
23:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Query Match	Length	ID	Description
1	24	100.0	5	22 AAB82036 Human hepreceptor
2	24	100.0	8	22 AAC68025 Human breast cancer
3	24	100.0	13	22 AAB82037 Human hepreceptor
4	24	100.0	29	22 ABB31539 Peptide #4198 enco
5	24	100.0	29	22 ABB36756 Peptide #4262 enco
6	24	100.0	29	22 ABB22083 Protein #4082 enco
7	24	100.0	29	22 AAM57583 Human brain expres
8	24	100.0	29	22 AAM69987 Human bone marrow
9	24	100.0	29	22 AAM17730 Peptide #4164 enco
10	24	100.0	29	22 AAM30245 Peptide #4282 enco

11	24	100.0	29	22 AAM05392 Peptide #4074 enco
12	24	100.0	29	23 AAG39536 Human hepreceptor
13	24	100.0	34	22 AAB82020 Novel human secret
14	24	100.0	52	22 AAU33060 Arabidopsis thalia
15	24	100.0	56	21 AAG60550 Human prostate tum
16	24	100.0	56	20 AAY74237 Arabidopsis thalia
17	24	100.0	63	21 AAG60549 Human cardiovascular
18	24	100.0	68	22 AAU21878 Antigen 2 from cl
19	24	100.0	84	20 AAB89863 Soybean Mito homolo
20	24	100.0	104	21 AAB91796 Human polypeptide
21	24	100.0	106	22 AAO01456 Human GEF protein
22	24	100.0	109	23 ABP08192 Amino acid sequenc
23	24	100.0	117	20 AAY75233 Listeria monocytog
24	24	100.0	121	23 ABB49483 Arabidopsis thalia
25	24	100.0	160	21 AAG59750 Arabidopsis thalia
26	24	100.0	164	21 AAG59749 Arabidopsis thalia
27	24	100.0	164	22 AAU35818 Helicobacter pylor
28	24	100.0	164	22 AAT45940 Helicobacter pylor
29	24	100.0	165	21 AAG03876 Human secreted pro
30	24	100.0	165	22 AAM40267 Human polypeptide
31	24	100.0	166	21 AAG59748 Arabidopsis thalia
32	24	100.0	168	13 AAR24252 Vaccinia virus sal
33	24	100.0	170	17 AAR98756 Murine NEBD-2 prot
34	24	100.0	170	17 AAR98756 Murine NEBD-2 prot
35	24	100.0	171	15 AAR45264 Murine NEBD-2. Mu
36	24	100.0	171	15 AAR45273 Murine NEBD-2 muta
37	24	100.0	171	15 AAR45274 Murine NEBD-2 muta
38	24	100.0	171	15 AAR45275 Murine NEBD-2 muta
39	24	100.0	171	21 AAR14251 Mouse nedd-2 prote
40	24	100.0	184	22 AAM42053 Human polypeptide
41	24	100.0	211	21 AAG11284 Arabidopsis thalia
42	24	100.0	222	21 AAG57892 Arabidopsis thalia
43	24	100.0	222	21 AAG51044 Arabidopsis thalia
44	24	100.0	222	21 AAG61659 Arabidopsis thalia
45	24	100.0	230	21 AAG11263 Arabidopsis thalia

ALIGNMENTS

RESULT 1				
AAB82036 standard; peptide: 5 AA.				
ID	AAB82036			
XX	AC	AAB82036		
XX	AC	AAB82036		
DE	13-JUN-2001	(first entry)		
XX	Human hepreceptor domain A binding peptide Eper24.			
XX	Human hepreceptor, cytostatic, anti-HIV, antidiabetic,			
KW	neutrophic, immune response inducer, ezrin, infectious diseases, cancer;			
KW	HIV related dementia.			
XX	Homo sapiens.			
XX	GB2354241-A.			
XX	GB2354241-A.			
PD	21-MAR-2001.			
XX	17-SEP-1999;	99GB-0021881.		
XX	17-SEP-1999;	99GB-0021881.		
XX	(HOLM/) HOLMS K D.			
XX	HOLMS RD,			
XX	WFI, 2001-293287/31.			
XX	Novel regulatory or unfolding peptides of ezrin that binds to			
PT	hepreceptor, useful for inducing immune response for treating			
PT	infectious diseases and cancer			



Db 1 KEELM 5

## RESULT 4

ABB31539  
ID ABB31539 standard; Peptide: 29 AA.

XX AC ABB31539;

XX DT 01-FEB-2002 (first entry)

XX DE Peptide #4190 encoded by breast cell single exon nucleic acid probe.

XX KW Human; microarray, single exon probe, gene expression; breast;

XX OS disease; cancer.

XX XX Homo sapiens

XX PN WO200157271-A2.

XX PD 09-AUG-2001

XX PF 30-JAN-2001; 2001WO-050662

XX PR 04-FEB-2000; 2000US-0180412

XX PR 26-MAY-2000; 2000US-0207456

XX PR 30-JUN-2000; 2000US-0608408

XX PR 03-AUG-2000; 2000US-0632466

XX PR 21-SEP-2000; 2000US-0234687

XX PR 27-SEP-2000; 2000US-0236359

XX PR 04-OCT-2000; 2000US-0024263

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001 496933/54.

XX PT New spatially addressable set of single exon nucleic acid probes,

XX PT useful for measuring gene expression in sample derived from human

XX PT breast, comprises number of single exon nucleic acid probes

XX PS Claim 27; SEQ ID NO 11507; 327pp + sequence listing; English.

XX CC The invention relates to a spatially addressable set of single exon

XX CC nucleic acid probes for measuring gene expression in a sample derived

XX CC from human breast and PI 474 cells. The method involves contacting

XX CC the probes with a collection of detectably labeled nucleic acids

XX CC derived from mRNA of human breast, and then measuring the label

XX CC bound to each probe of the microarray. The probes are useful for

XX CC verifying the expression of regions of genomic DNA predicted to

XX CC encode proteins. They are useful for gene discovery, and for

XX CC determining predisposition and/or prognosing breast disease. Gene

XX CC expression analysis is useful for assessing the toxicity of chemical

XX CC agents on cells. The microarray of this invention presents a far greater

XX CC diversity of probes for measuring gene expression, with far less bias

XX CC than expressed sequence tag microarrays. The method is suitable for

XX CC rapid production of functional information from genomic sequence. The

XX CC present sequence is a peptide encoded by a single exon nucleic acid

XX CC probe of the invention.

XX CC Note: The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 29 AA:

Query Match 100.0%; Score 24; DB 22; Length 29;

Best Local Similarity 100.0%; Pred. No. 1e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEELM 5

|||||

Db 8 KEELM 12

## RESULT 5

ABB36756  
ID ABB36756 standard; Peptide: 29 AA.

XX AC ABB36756;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #4262 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver, gene expression, single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-050662

XX PR 04-FEB-2000; 2000US-0180312

XX PR 26-MAY-2000; 2000US-0207456

XX PR 30-JUN-2000; 2000US-0608408

XX PR 03-AUG-2000; 2000US-0632366

XX PR 21-SEP-2000; 2000US-0234687

XX PR 27-SEP-2000; 2000US-0236359

XX PR 04-OCT-2000; 2000US-0024263

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001 483447/52.

XX PT Human genome derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human fetal liver

XX PS Claim 27; SEQ ID NO 29391; 639pp + sequence listing; English.

XX CC The invention relates to a single exon nucleic acid probe for

XX CC measuring human gene expression in a sample derived from human foetal

XX CC liver. The single exon nucleic acid probes may be used for predicting,

XX CC measuring and displaying gene expression in samples derived from human

XX CC foetal liver. The present sequence is a peptide encoded by a single exon

XX CC nucleic acid probe of the invention.

XX CC Note: The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 29 AA:

Query Match 100.0%; Score 24; DB 22; Length 29;

Best Local Similarity 100.0%; Pred. No. 1e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEELM 5

|||||

Db 8 KEELM 12

## RESULT 6

ABB22083  
ID ABB22083 standard; Protein: 29 AA.

XX AC ABB22083;

XX DT 23 JAN 2002 (first entry)

XX DE Protein #4082 encoded by probe for measuring heart cell gene expression.

XX KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.

OS Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001

XX 40-JAN-2001; 2001WO-US00666

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0642366.

XX 21-SEP-2000; 2000US-0344687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000US-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human hearts.

XX Claim 15; SEQ ID No 24854; 540pp; English.

XX The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see AB21545-AM41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays.

XX By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.

XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 29 AA;

XX Query Match 100.0%; Score 24; DB 22; Length 29;

XX Best Local Similarity 100.0%; Pred. No. 1e+02;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 KEELM 5

XX 8 KEELM 12

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0642366.

XX 21-SEP-2000; 2000US-0344687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000US-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human brains.

XX Example 4; SEQ ID No: 29608; 650pp; English.

XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.

XX Sequence 29 AA;

XX Query Match 100.0%; Score 24; DB 22; Length 29;

XX Best Local Similarity 100.0%; Pred. No. 1e+02;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 KEELM 5

XX 8 KEELM 12

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WPI: 2001-488900/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4: SEQ ID NO: 30213; 658pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 29 AA:  
 Query Match 100.0%; Score 24; DB 22; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KEELM 5  
 DB 8 KEELM 12  
 RESULT 9  
 AAM17730  
 ID AAM17730 standard; Protein; 29 AA.  
 AC AAM17730;  
 DE 12-OCT-2001 (first entry)  
 PE Peptide #4164 encoded by probe for measuring cervical gene expression.  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 KW Homo sapiens  
 OS Homo sapiens  
 PN WC230;57778-A2  
 XX 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00670  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI: 2001 488901/53  
 DP Human genome derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID NO 22556, 487pp; English.  
 XX The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see AAM10068-AAM28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably

CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 29 AA:  
 Query Match 100.0%; Score 24; DB 22; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KEELM 5  
 DB 8 KEELM 12  
 RESULT 10  
 AAM30245  
 ID AAM30245 standard; Protein; 29 AA.  
 XX AAM30245;  
 AC AAM30245;  
 DE 17-OCT-2001 (first entry)  
 PE Peptide #4282 encoded by probe for measuring placental gene expression.  
 DE Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN W0200157272-A2.  
 XX 09-AUG-2001  
 PD 30-JAN-2001; 2001WO-US00663.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI: 2001-488997/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27; SEQ ID No 30514; 654pp; English.  
 XX The present invention relates to single exon nucleic acid probes (SENP;  
 CC see AAM31315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC prediction, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 XX  
 SQ Sequence 29 AA:  
 Query Match 100.0%; Score 24; DB 22; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KEELM 5  
 DB 8 KEELM 12

```

RESULT 11
AAM05492
ID AAM05492 standard; Protein; 29 AA.
XX
AC AAM05492;
XX
DI 09-OCT-2001 (first entry)
DE Peptide #4074 encoded by probe for measuring breast gene expression
XX
XX Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN W0200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29 JAN 2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180412
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234487.
PR 27-SEP-2000; 2000US-0246459.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WI: 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PI in a human breast.
XX
PS claim 27; SEQ ID NO 14132; 422pp; English.
XX
XX the present invention relates to novel single exon nucleic acid probes
CC (see AAM00010-AA10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer; disorders of development,
CC inflammatory diseases of the breast; fibrocystic changes; Proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published\_pct\_sequences.
XX
SU Sequence 29 AA;
Query Match 100.0%; Score 24; DB 22; Length 29,
Best Local Similarity 100.0%; Pred. No. 1e-02;
Matches 5; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
DY 1 KEELM 5
DI 8 KEELM 12
RESULT 12
AAG39536
ID AAG39536 standard; Peptide; 29 AA.
XX
AC AAG39536;
XX
DI 19-AUG-2002 (first entry)

```

```

XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 29201.
XX
XX Human, single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Helmsky-Pudlak syndrome; sarcoidosis; pulmonary haemangiomas;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN W0200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180412P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-246359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WI: 2002-114183/15.
XX
XX Spatially addressable set of single exon nucleic acid probes, used to
PI measure gene expression in human lung samples.
XX
PS claim 27; SEQ ID NO 29201; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung, measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Helmsky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemangiomas, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

```

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences.  
 XX  
 SQ Sequence 29 AA;

Query Match 100.0%; Score 24; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 KEELM 5  
 DB 8 KEELM 12  
 I I I I I

RESULT 13  
 ID AAB82020 standard, peptide, 34 AA.  
 AC AAB82020;  
 DE 13-JUN-2001 (first entry)  
 KW Human hepreceptor domain B  
 KW Human; hepreceptor domain B; cytotostatic; anti-HIV; antihistonic;  
 KW neurotrophic; immune response inducer; ezrin; infectious diseases; cancer;  
 KW HIV-related dementia;  
 OS Homo sapiens.  
 PI Key location/Qualifiers  
 FI Modified-site 14 /note= "Optionally phosphorylated"  
 XX GB2354241-A.  
 XX 21-MAR-2001.  
 XX 17-SEP-1999; 99CH-0021881.  
 XX 17-SEP-1999; 99GB-0021881.  
 XX (HOLM/) HOLMS P D.  
 PI Holms RD;  
 XX WPI: 2001 203287/31.  
 PI Novel regulatory or unfolding peptides of ezrin that binds to  
 PI hepreceptor, useful for inducing immune response for treating  
 PI infectious diseases and cancer.  
 PS Claim 5, Page 36; 42pp. English  
 XX The present sequence is domain B of human hepreceptor of human ezrin. The  
 CC hepreceptor is a novel active site in human ezrin. Ezrin regulates the  
 CC structure of the cortical cytoskeleton to control cell surface  
 CC topography. The present invention relates to peptides (see AAB82041 to  
 CC AAB82044) that bind to hepreceptor with greater affinity than HEP1 (see  
 CC AAB82046). The hepreceptor binding peptides are useful for inducing  
 CC immune response, and for treating infectious diseases, cancer and  
 CC HIV-related dementia. The present sequence assembles into two  
 CC anti-parallel helices with hepreceptor domain A (see AAB82019).  
 XX Sequence 34 AA;

Query Match 100.0%; Score 24; DB 22; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02.

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEELM 5  
 DB 4 KEELM 8  
 I I I I I

RESULT 14  
 ID AAB33060 standard; Protein; 52 AA.  
 XX AAB33060;  
 AC AAB33060;  
 DE 18-DEC-2001 (first entry)  
 KW Novel human secreted protein #3551.  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 OS Homo sapiens.  
 PI WO200179449-A2.  
 XX 25-OCT-2001.  
 XX 16-APR-2001; 2001WO-US08656.  
 XX 18-APR-2000; 2000US-0552929.  
 XX 25-JAN-2001; 2001US-073150.  
 XX (HYSE-) HYSEQ INC.  
 PI Tanq VT, Liu C, Drmanac RT;  
 FI WPI: 2001-611725/70.  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 XX vaccination, testing and therapy.  
 PS Claim 20; Page 702; 765pp; English.  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising  
 CC the nucleic acids encoding the polypeptides and cells genetically  
 CC engineered to express them are also useful for producing the proteins.  
 CC The proteins are useful in genetic vaccination, testing and  
 CC therapy, and can be used as nutritional supplements. They may be used to  
 CC increase stem cell proliferation; to regulate haematopoiesis; and in  
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
 CC in treatment of leukaemia. AAB29510-AAB33060 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.  
 XX Sequence 52 AA;

Query Match 100.0%; Score 24; DB 22; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEELM 5  
 DB 12 KEELM 16  
 I I I I I

RESULT 15

AA60550  
ID AA60550 standard; Protein; 56 AA.  
XX  
XX AC AA60550;  
XX  
XX DT 18 OCT 2000 (first entry)  
XX  
XX DE  
XX  
XX KW Arabidopsis thaliana protein fragment SEQ ID NO: 78439.  
XX  
XX KW Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX OS Arabidopsis thaliana.  
XX  
XX PN EPI033405-A2.  
XX  
XX PD 06 SEP 2000.  
XX  
XX PF 25-FEB-2000; 2000EP 0301439.  
XX  
XX PF 25-FEB-1999; 990S-0121825.  
XX PF 05-MAR-1999; 990S-0124180.  
XX PF 09-MAR-1999; 990S-0124548.  
XX PF 23-MAR-1999; 990S-0125788.  
XX PF 25-MAR-1999; 990S-0126264.  
XX PF 29-MAR-1999; 990S-0126785.  
XX PF 01-APR-1999; 990S-0127462.  
XX PF 06-APR-1999; 990S-0128234.  
XX PF 08-APR-1999; 990S-0128714.  
XX PF 16-APR-1999; 990S-0129845.  
XX PF 19-APR-1999; 990S-0130077.  
XX PF 21-APR-1999; 990S-0130449.  
XX PF 23-APR-1999; 990S-0130510.  
XX PF 23-APR-1999; 990S-0130891.  
XX PF 28-APR-1999; 990S-0131449.  
XX PF 30-APR-1999; 990S-0132048.  
XX PF 30-APR-1999; 990S-0132407.  
XX PF 04-MAY-1999; 990S-0132484.  
XX PF 05-MAY-1999; 990S-0132485.  
XX PF 06-MAY-1999; 990S-0132486.  
XX PF 07-MAY-1999; 990S-0132487.  
XX PF 11-MAY-1999; 990S-0132863.  
XX PF 14-MAY-1999; 990S-0134256.  
XX PF 14-MAY-1999; 990S-0134218.  
XX PF 14-MAY-1999; 990S-0134219.  
XX PF 14-MAY-1999; 990S-0134221.  
XX PF 14-MAY-1999; 990S-0134370.  
XX PF 16-MAY-1999; 990S-0134768.  
XX PF 19-MAY-1999; 990S-0134941.  
XX PF 20-MAY-1999; 990S-0135124.  
XX PF 21-MAY-1999; 990S-0135353.  
XX PF 24-MAY-1999; 990S-0135629.  
XX PF 25-MAY-1999; 990S-0136021.  
XX PF 27-MAY-1999; 990S-0136392.  
XX PF 28-MAY-1999; 990S-0136782.  
XX PF 01-JUN-1999; 990S-0137222.  
XX PF 03-JUN-1999; 990S-0137528.  
XX PF 04-JUN-1999; 990S-0137502.  
XX PF 07-JUN-1999; 990S-0137724.  
XX PF 08-JUN-1999; 990S-0138094.  
XX PF 10-JUN-1999; 990S-0138540.  
XX PF 10-JUN-1999; 990S-0138847.  
XX PF 14-JUN-1999; 990S-0139119.  
XX PF 16-JUN-1999; 990S-0139452.  
XX PF 16-JUN-1999; 990S-0139453.  
XX PF 17-JUN-1999; 990S-0139492.  
XX PF 18-JUN-1999; 990S-0139454.  
XX PF 18-JUN-1999; 990S-0139455.  
XX PF 18-JUN-1999; 990S-0139456.  
XX PF 18-JUN-1999; 990S-0139458.  
XX PF 18-JUN-1999; 990S-0139458.  
XX PF 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142555.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144384.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149468.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PP



PR 23-AUG-1999; 99US-0149902.  
PR 24-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158469.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 24; DB 21; Length 56;  
Best Local Similarity 100.0%; Pred. No. 20-02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEELM 5  
DB 10 KEELM 14

Search completed: January 16, 2003, 16:49:13  
Job time : 24.7857 secs

